Implied non-recombining trees and calibration for the volatility smile

Chris Charalambous[1*], Nicos Christofides[2], Eleni D. Constantinide[1], and Spiros H. Martzoukos[1]

May 2007

JEL classification: C61, C63, G13, G15

Keywords: Volatility Smile, Implied Non-recombining Trees, Calibration, Non-Convex Constrained Optimization.

Acknowledgements: The authors are grateful for financial support from a research grant on Contingent Claims from the University of Cypus.

---

[1] University of Cyprus.
[2] Centre for Quantitative Finance, Imperial College, London.
**Corresponding Author*
Charalambous C., Professor
Department of Business Administration, University of Cyprus
P.O. Box 20537, CY 1678 Nicosia, Cyprus
Tel.: +357-22892466, Fax: +357-22892460
e-mail: bachris@ucy.ac.cy

Abstract

In this paper we capture the implied distribution from option market data using a non-recombining (binary) tree, allowing the local volatility to be a function of the underlying asset and of time. The problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure and use nonlinear constrained optimization to minimize the least squares error function on market prices. The proposed model can accommodate European options with single maturities and, with minor modifications, options with multiple maturities. It can provide a market-consistent tree for option replication with transaction costs (often this requires a non-recombining tree) and can help pricing of exotic and Over The Counter (OTC) options. We test our model using options data of the FTSE 100 index obtained from LIFFE. The results strongly support our modelling approach.

I. Introduction

Calibrating a tree, otherwise known as constructing an implied tree, means finding the stock price and/or associated probability at each node in such a way that the tree reproduces the current market prices for a set of benchmark instruments. The main benefit of calibrating a model to a set of observed option prices is that the calibrated model is consistent with today's market prices. The calibrated model can then be used to price other more complex or less liquid securities, such as (OTC) options whose prices may not be available in the market.

The binomial tree is the most widely used tool in the financial pricing industry. The classic Cox-Ross-Rubinstein (CRR, 1979) binomial tree is a discretization of the Black-Scholes (BS, 1973) model since it is based on the assumption of the BS model that the underlying asset evolves according to a geometric Brownian motion with a constant volatility factor. This, however, contradicts the observed implied volatility, which suggests that volatility depends on both the strike and maturity of an option, a relationship commonly known as the volatility smile. This problem has motivated the recent literature on "*smile consistent*" *no-arbitrage models*. Consistency is achieved by extracting an implied evolution for the stock price from market prices of liquid standard options on the underlying asset. There are two classes of methodologies within this approach. Smile consistent *deterministic volatility* models (Rubinstein, 1994, Derman and Kani, 1994, Dupire, 1994, Barle and Kakici, 1995, Rubinstein and Jackwerth, 1996, Jackwerth, 1997, etc.); and *stochastic volatility* smile consistent models which allow for smile-consistent option pricing under the no-arbitrage evolution of the volatility surface (Derman and Kani, 1998, Ledoit and Santa-Clara, 1998, Britten-Jones and Neuberger, 2000, etc.). The latter class of models is more general and it nests the former class of models (Skiadopoulos, 2001). There also exist *non-parametric* methods, like Stutzer (1996) who uses the maximum entropy concept to derive the risk neutral distribution from the historical distribution of the asset price and Ait-Sahalia and Lo (1998) who propose a non-parametric estimation procedure for state-price densities using observed option prices.

Smile consistent deterministic volatility models are based on the assumption that the local volatility of the underlying asset is a known function of time and of the path and level of the underlying asset price. However, they do not specify local volatility in advance, but derive it endogenously from the European option prices. Therefore, they preserve the "pricing by no-arbitrage" property of the BS model, and the markets are complete since the option's pay-off can be synthesized from existing assets.

Rubinstein (1994) finds the implied risk-neutral terminal-node probability distribution which is in the least-squares sense, closest to the lognormal subject to some constraints. The probabilities must add up to one and be non-negative. Moreover, they are calculated so that the present value of the underlying assets and all the European options calculated with these probabilities, fall between their respective bid-ask prices. This methodology allows for an arbitrary terminal-node probability distribution, but assumes that path probabilities leading to the same ending node are equal. Rubinstein's (1994) methodology suffers from the fact that options expiring at early time steps cannot be used for the construction of the tree. Thus, options with maturity other than the maturity of the options used during the construction of the tree are not consistent with market prices.

Jackwerth (1997) introduced generalized binomial trees as an extension of Rubinstein (1994). His model allows for an arbitrary terminal-node probability distribution, but also allows path probabilities leading to the same node to take different values.

Derman and Kani (1994) and Dupire (1994) constructed recombining binomial trees using a large set of option prices. For each node they need a corresponding option price with strike price equal to the node's stock price and expiring at the time associated with that node. Since they have fewer option prices than required, they need to interpolate and extrapolate from given option prices. Their trees are sensitive to the interpolation and extrapolation method and require adjustments to avoid arbitrage violations.

Barle and Cakici (1995) introduced a number of modifications which aimed to eliminate negative probabilities and improve the general stability of Derman's and Kani's (1994) model. Although their modified method fits the smile accurately, negative probabilities may still occur with increases in the volatility smile and interest rate. As they state, this is because of their "… strict requirement that continuous diffusion be modeled as a binomial process and on a recombining tree ". This problem can be referred to as a problem of *interdependencies* between nodes.

Possible methods that can be used to reduce the problem of inter-dependencies are the calibration of trinomial (or multinomial) trees or non-recombining trees. These extra degrees of freedom allow for more flexibility in the estimation of the distribution of the underlying asset.

Trinomial trees provide a much better approximation to the continuous time process than the binomial trees for the same number of steps. However, the extra degrees of freedom (additional number of nodes) require a larger number of simultaneous equations to be solved. Derman, Kani and Chriss (1996) proposed implied trinomial trees. In their model they use the additional parameters to conveniently choose the "state space" of all node prices in the tree, and let only the transition probabilities be constrained by market options prices. Chriss (1996) generalized their method for American style options.

In this paper we propose a method for calibrating a non-recombining (binary) tree, based on optimization. Specifically, we minimize the discrepancy between the observed market prices and the theoretical values with respect to the underlying asset at each node, subject to constraints that maintain risk neutrality and prevent arbitrage opportunities. Our model is built on a non-recombining tree[1] so as to allow the local volatility to be a

[1] Other work we are aware of that uses a non-recombining tree is of Talias (2005) where for the calibration he uses genetic algorithms.

function of the underlying asset and of time and to enable each node of the tree to act as an independent variable. Effectively, the problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure, and using methods from nonlinear constrained optimization we minimize the least squares error function. Specifically, we adopt a penalty method and for the optimization we use a Quasi-Newton algorithm. Because of the combinatorial nature of the tree and the large number of constraints, the search for an optimum solution as well as the choice of an algorithm that performs well becomes a very challenging problem.

Our model was created as a response for the need of a non-recombining implied tree. The main benefit of the model is its *analytical structure* which enables us to use efficient methods for nonlinear optimization. Although the method uses a large number of variables, due to the fact that we use efficient methods for optimization the model is not computationally intensive. Also, the proposed methodology can be easily modified to capture the observed bid/ask spreads in the market. This is very useful since the reported closing prices may not always be accurate, or may be inaccurate due to various market frictions. In addition, calibration of the non-recombining tree can be used for option replication with transaction costs as in Edirisinghe et al. (1993) and other related methodologies that require non-recombining trees.

In contrast to Rubinstein (1994), the proposed methodology can be easily modified to account for European contracts with different maturities. Our method does not need any interpolation or extrapolation across strikes and time to find hypothetical options as opposed to Derman and Kani (1994). Finally, the extra degrees of freedom and the analytical structure of the model would allow us to impose smoothness constraints on the distribution of the underlying asset if required.

We test our model using options data on the FTSE 100 index, for the year 2003 obtained from LIFFE. The results strongly support our modelling approach. Pricing results are smooth without the presence of an over-fitting problem and the derived implied distributions are realistic. Also, the computational burden is not a major issue.

The paper continues as follows: In section II we describe the proposed methodology and the initialization of the non-recombining tree. In section III we discuss the imposed risk neutrality and no-arbitrage constraints. In section IV we describe the optimization algorithm. In section V we test the model using FTSE 100 options data. Conclusions are in section VI. In Appendix 1A we prove the feasibility of the initialized tree, in Appendix 1B we prove the feasibility of the initialized tree taking into account that the risk-free rate, dividend yield and time step are time dependent and in Appendix 2 we adjust the formulas for time dependent risk free rate, dividend yield and step size.

II. The proposed methodology and initialization of the non-recombining tree

Our goal is to develop an arbitrage-free risk neutral model that fits the smile, is preference-free, and can be used to value options from easily observable data. In order to allow more degrees of freedom, we use a non-recombining tree. In the following section we present the proposed methodology, and describe the initialization of the tree.

Figure 1 shows a non-recombining tree with four steps.

[Insert figure 1 here]

The point (i, j) on the tree denotes:

i : the time dimension, $i = 1, ..., n$

j : the asset (time specific) dimension, $j = 1, ..., 2^{i-1}$

$S(i, j)$ is the value of the underlying asset at node (i, j).

Figure 2 shows a typical triplet in a non-recombining tree.

[Insert figure 2 here]

Let $C_{Mkt}(k)$, $k=1,...,N$ denote the market prices of N European calls, with strikes $K(k)$ and single maturities *T*. Also, let $C_{Mod}(x,k)$, $k=1,...,N$ denote the theoretical prices of the N calls obtained using the model. x denotes a vector containing the variables of the model which are the values of the underlying asset at each node of the tree, excluding its current value. The ideal solution is to find the values of the underlying asset (the model variables) at each node of the tree such that a perfect match is achieved between the option market prices and those predicted by the tree. However, due to market imperfections and other factors perfect matching may not always be possible. Therefore, we minimize the discrepancy between the observed market prices and the theoretical values produced by the model subject to constraints that prevent arbitrage opportunities.

We have to solve a non-convex constrained minimization problem with respect to the values of the underlying asset at each node:

$$\min_{x} \sum_{k=1}^{N} w_k f\left(C_{Mod}(x,k), C_{Mkt}(k)\right) \quad (1)$$

where f denotes a suitable objective function on the error between the observed and market prices. We can also allow for a weight factor, w_k[2]. In this paper we use the least squares error function which is defined as the sum of square differences between market prices and theoretical prices produced by the tree. The method can be adjusted easily for any other objective function.

The philosophy of the initialization of the non-recombining tree is the same as that of the construction of the standard CRR binomial tree, but we adjust the formulas so that the tree does not necessarily recombine.

We denote with $u(i,j)$ and $d(i,j)$ the up and down factors by which the underlying asset price can move in the single time step, Δt, given that we are at node (i,j). Δt, $u(i,j)$ and $d(i,j)$ factors are given by the following formulas[3]:

---

[2] Weights can be related for example to the trading volume of the options.

[3] For simplicity, we make the assumption that the risk free rate, the dividend yield and the step size do not change across time. Formulas adjusted for time dependence can be found in Appendix 1B and 2.

$$\Delta t = \frac{T}{n-1} \tag{2}$$

$$u(i,j) = e^{\sigma(i)\sqrt{\Delta t}} \tag{3a}$$

$$i = 1,\ \ldots,\ n-1\ ,\quad j = 1,\ \ldots,\ 2^{i-1}$$

$$d(i,j) = e^{-\sigma(i)\sqrt{\Delta t}} = \frac{1}{u(i,j)} \tag{3b}$$

where T is the option's time to maturity and $\sigma(i)$ is the volatility term structure at time step i.

We initialize the tree using the following volatility term structure:

$$\sigma(i) = \sigma(1)e^{\lambda(i-1)\Delta t}\ ,\ \lambda \in R\ ,\ i = 1,\ \ldots,\ n\text{-}1 \tag{4}$$

where λ is a constant parameter and $\sigma(1)$ is a properly chosen initial value for the volatility. If λ is positive, then volatility increases as we approach maturity and if λ is negative, then volatility decreases as we approach maturity[4].

In order to preserve the risk neutrality at every time step and hence obtain a feasible initial tree, we choose λ to belong in the following interval (for proof see Appendix 1A):

$$\lambda \in \left[\frac{1}{T}\log\left(\frac{|r_f - \delta|\sqrt{\Delta t}}{\sigma(1)} \right),\quad +\infty \right) \tag{5}$$

By choosing λ from the above interval, we allow the initial volatility to increase or decrease across time. We make several consecutive draws from interval (5) until we find the value of λ that gives the "optimal" tree[5].

We denote with $S(1,1)$ the current value of the underlying asset. The odd nodes of the tree $S(i,j)$, are initialized using the following equation:

$$S(i,j) = S(i-1,\frac{j+1}{2})d(i-1,\frac{j+1}{2})\ ,\ i = 2, \ldots, n,\ \ j = 1, 3, \ldots, 2^{i-1}-1 \tag{6a}$$

The even nodes of the tree $S(i,j)$, are initialized using the following equation:

---

[4] Other non-monotonic functions could also be used for $\sigma(i)$ but what we have tried proved adequate for our purposes.

[5] Optimal tree is the one that gives the lowest-value objective function subject to the initial constraints.

$$S(i,j)=S(i-1,\frac{j}{2})u(i-1,\frac{j}{2}),\quad i=2,\dots,n,\quad j=2,4,\dots,2^{i-1} \tag{6b}$$

We want to point out that equations (3) to (6) are used only for initialization. Once the optimization process starts, each value of the underlying asset (except from $S(1,1)$) acts as an independent variable in the system.

Upward transition probabilities give the probability of moving from node (i,j) to node $(i+1,2j)$ whereas downward transition probabilities give the probability of moving from node (i,j) to node $(i+1,2j-1)$ for $i=1,\dots,n-1$ and $j=1,\dots,2^{i-1}$. For the upward transition probabilities $p(i,j)$ between the various nodes of the tree we use the risk-neutral probability formula:

$$p(i,j)=\frac{S(i,j)e^{(r_f-\delta)\Delta t}-S(i+1,2j-1)}{S(i+1,2j)-S(i+1,2j-1)},\quad i=1,\dots,n-1,\ \ j=1,\dots,2^{i-1} \tag{7}$$

where r_f denotes the annually continuously compounded riskless rate of interest and δ denotes the annually continuously compounded dividend yield. Their respective downward probability is equal to one minus the upward probability.

The call option value at the last time step is given by:

$$C(n,j)=\max\{S(n,j)-K,0\},\quad j=1,\dots,2^{n-1} \tag{8}$$

However, the function *max* is non differentiable at $S(n,j)=K$. To overcome this problem, we propose the following smoothing approximation to $C(n,j)$:

$$\frac{C_\alpha(n,j)}{K}=\begin{cases}0 & \textit{for } S(n,j)/K\le 1-z/2\\[2ex] \dfrac{S(n,j)}{K}-1 & \textit{for } S(n,j)/K\ge 1+z/2\\[2ex] \dfrac{1}{2z}\left[\left(\dfrac{S(n,j)}{K}-1\right)+\dfrac{z}{2}\right]^2 & \textit{for } 1-z/2<S(n,j)/K<1+z/2\end{cases} \tag{9a}$$

$$j=1,\dots,2^{n-1}$$

where z is a small positive constant, for example 0.01 (see Fig. 3).

[Insert figure 3 here]

The value of the call at intermediate nodes is given by the following equation:

$$C(i,j) = \left(p(i,j)C(i+1,2j) + (1-p(i,j))C(i+1,2j-1)\right)e^{-r_f \Delta t} \qquad (9b)$$

$$i = n-1, \ldots, 1, \quad j = 1, \ldots, 2^{i-1}$$

III. Risk neutrality and no-arbitrage constraints

In this section we describe the risk neutrality and no-arbitrage constraints. In order for the transition probabilities $p(i,j)$ defined in Eq.(7) to be well specified, they should take values between zero and one. This implies the following *risk-neutrality* constraints:

$$S(i,j)e^{(r_f-\delta)\Delta t} \leq S(i+1,2j) \qquad (10a)$$

$$i = 1, \ldots, n-1, \; j = 1, \ldots, 2^{i-1}$$

$$S(i,j)e^{(r_f-\delta)\Delta t} \geq S(i+1,2j-1) \qquad (10b)$$

Risk neutrality constraints in the non-recombining tree prevent nodes $2j-1$ and $2j$ to cross, for $i = 1,\ldots,n$ and $j = 1,\ldots,2^{i-1}$ (see Fig.1).

Options (puts and calls) have upper and lower bounds that do not depend on any particular assumptions on the factors that affect option prices. If the option price is above the upper bound or below the lower bound, there are profitable opportunities for arbitrageurs. To avoid such opportunities, we include the *no-arbitrage* constraints. Specifically, a European call with dividends should lie between the following bounds:

$$\max\left(S(1,1)e^{-\delta T} - Ke^{-r_f T}, 0\right) \leq C_{Mod} \leq S(1,1) \qquad (11)$$

Also, every value of the underlying asset on the tree should be greater or equal to zero. Thus, we also impose the following constraint:

$$S(i,j) \geq 0, \; i = 2, \ldots, n, \; j = 1, \ldots, 2^{i-1} \qquad (12)$$

IV. The optimization algorithm

The objective of the problem is to minimize the least squares error function of the discrepancy between the observed market prices and the theoretical values produced by the model. Thus, we have the following optimization problem:

$$\min_x \frac{1}{2}\sum_{k=1}^{N}\left(C_{Mod}(x,k)-C_{Mkt}(k)\right)^2 \tag{13}$$

where $C_{Mod}(k)$ and $C_{Mkt}(k)$ denote the model and market price respectively of the k^{th} call, $k=1,\ldots,N$, subject to the constraints:

i) $g_1(i,j)=S(i,j)e^{(r_f-\delta)\Delta t}-S(i+1,2j-1)\geq 0$, $\quad i=1,\ldots,n-1$, $j=1,\ldots,2^{i-1}$ (14a)

ii) $g_2(i,j)=S(i+1,2j)-S(i,j)e^{(r_f-\delta)\Delta t}\geq 0$, $i=1,\ldots,n-1$, $j=1,\ldots,2^{i-1}$ (14b)

iii) $g_3(k)=S(1,1)-C_{Mod}(k)\geq 0$, $k=1,\ldots,N$ (14c)

iv) $g_4(k)=C_{Mod}(k)-\max(S(1,1)e^{-\delta T}-K(k)e^{-r_f T},0)\geq 0$, $k=1,\ldots,N$ (14d)

v) $g_5(i,j)=S(i,j)\geq 0$, $i=2,\ldots,n$, $j=1,\ldots,2^{i-1}$ (14e)

Since the problem under consideration is a non-convex optimization problem with linear constraints we adopt an exterior penalty method (Fiacco and McCornick, 1968) to convert the nonlinear constrained problem into a nonlinear unconstrained problem. The Exterior Penalty Objective function that we use is the following:

$$\begin{aligned}P(x,\alpha)=&\frac{1}{2}\sum_{k=1}^{N}\left(C_{Mod}(x,k)-C_{Mkt}(k)\right)^2\\&+\frac{\alpha}{2}\sum_{i=1}^{n-1}\sum_{j=1}^{2^{i-1}}\left(\left[\min\left(g_1(i,j),0\right)\right]^2+\left[\min\left(g_2(i,j),0\right)\right]^2\right)\\&+\frac{\alpha}{2}\sum_{k=1}^{N}\left(\left[\min\left(g_3(k),0\right)\right]^2+\left[\min\left(g_4(k),0\right)\right]^2\right)\\&+\frac{\alpha}{2}\sum_{i=2}^{n}\sum_{j=1}^{2^{i-1}}\left(\left[\min\left(g_5(i,j),0\right)\right]^2\right)\end{aligned}\tag{15}$$

The second, third and fourth terms in $P(x,\alpha)$ give a positive contribution if and only if x is infeasible. Under mild conditions it can be proved that minimizing the above penalty function for strictly increasing sequence α tending to infinity the optimum point $x(\alpha)$ of P tends to x^*, a solution of the constrained problem.

For the optimization we use a Quasi-Newton algorithm. Specifically we use the BFGS formula[6] (Fletcher, 1987). For the procedure of *Line Search* in the algorithm we use the Charalambous (1992) method. To achieve the best feasible solution, i.e. the solution that gives us a feasible tree with the smallest error function we force the algorithm to draw consecutively values of λ from the specified interval (5) until the objective function is smaller than 1.E-4 and also the penalty term equals zero, i.e. we have a feasible solution.

Implementation

For the implementation of the optimization method, we need to calculate the partial derivatives of $C_{Mod}(k)$ [7] with respect to the value of the underlying asset at *each node*, for $k = 1, ..., N$ i.e. we want to find $\dfrac{\partial C(1,1,k)}{\partial S(i,j)}$, $i = 2, ..., n$, $j = 1, ..., 2^{i-1}$ [8] and $k = 1, ..., N$. For notational simplicity in the following, we assume that we have only one call option. For the computation of $\dfrac{\partial C(1,1)}{\partial S(i,j)}, \forall i, j$ we implement the following steps:

We define the triplet vector (see Fig.2):

$$S_{i,j}^{(l)} = [S(i,j) \;\; S(i+1,2j) \;\; S(i+1,2j-1)] \tag{16}$$

1^st^ step: Compute the partial derivatives of the risk neutral transition probabilities, $\dfrac{\partial p(i,j)}{\partial S(i,j)}$, $\dfrac{\partial p(i,j)}{\partial S(i+1,2j)}$ and $\dfrac{\partial p(i,j)}{\partial S(i+1,2j-1)}$ for $i = 1, ..., n-1$, and $j = 1, ..., 2^{i-1}$. We summarize the derivatives in vector form (17).

[6] The BFGS formula was discovered in 1970 independently by Broyden, Fletcher, Goldfarb and Shanno.
[7] From now on we will use $C(1,1)$ instead of C_{Mod}.
[8] We do not calculate $\dfrac{\partial C(1,1,k)}{\partial S(1,1)}$ since $S(1,1)$ is a known, fixed parameter, and thus does not take part in the optimization.

$$\nabla_{S_{i,j}^{(l)}} p(i,j) \equiv \begin{bmatrix} \partial p(i,j) \big/ \partial S(i,j) \\ \partial p(i,j) \big/ \partial S(i+1,2j) \\ \partial p(i,j) \big/ \partial S(i+1,2j-1) \end{bmatrix} = \frac{1}{S(i+1,2j) - S(i+1,2j-1)} \begin{bmatrix} e^{(r_f - \delta)\Delta t} \\ -p(i,j) \\ -(1-p(i,j)) \end{bmatrix} \tag{17}$$

2nd step: Compute the partial derivatives $\dfrac{\partial C(i,j)}{\partial S(i,j)}$, for $i = 2, ..., n-1$, and $j = 1, ..., 2^{i-1}$, $\dfrac{\partial C(i,j)}{\partial S(i+1,2j)}$ and $\dfrac{\partial C(i,j)}{\partial S(i+1,2j-1)}$ for $i = 1, ..., n-1$, $j = 1, ..., 2^{i-1}$.

We summarize the derivatives in vector form (18).

$$\nabla_{S_{i,j}^{(l)}} C(i,j) \equiv \begin{bmatrix} \partial C(i,j) \big/ \partial S(i,j) \\ \partial C(i,j) \big/ \partial S(i+1,2j) \\ \partial C(i,j) \big/ \partial S(i+1,2j-1) \end{bmatrix} = \begin{bmatrix} \Delta(i,j) \\ p(i,j)\left(\Delta(i+1,2j) - \Delta(i,j)e^{\delta\Delta t}\right)e^{-r_f\Delta t} \\ \left(1-p(i,j)\right)\left(\Delta(i+1,2j-1) - \Delta(i,j)e^{\delta\Delta t}\right)e^{-r_f\Delta t} \end{bmatrix} \tag{18}$$

where

$$\Delta(i,j) = \frac{C(i+1,2j) - C(i+1,2j-1)}{S(i+1,2j) - S(i+1,2j-1)}\, e^{-\delta\Delta t} = \frac{\partial C(i,j)}{\partial S(i,j)} \equiv Delta\ \ Ratio \tag{19}$$

3rd step: Compute the partial derivatives $\dfrac{\partial C_\alpha(n,j)}{\partial S(n,j)}$ for $j = 1, ..., 2^{n-1}$. They are given by the following formula:

$$\frac{\partial C_\alpha(n,j)}{\partial S(n,j)} = \begin{cases} 0 & for\ S(n,j) \le K(1 - z/2) \\ 1 & for\ S(n,j) \ge K(1 + z/2) \\ \dfrac{1}{z}\left[\left(\dfrac{S(n,j)}{K} - 1\right) + \dfrac{z}{2}\right] & for\ K(1 - z/2) < S(n,j) < K(1 + z/2) \end{cases} \tag{20}$$

4th step: Compute the partial derivatives $\frac{\partial C(1,1)}{\partial S(i,j)}$ for $i \geq 3$.

$$\frac{\partial C(1,1)}{\partial S(i,j)} = \prod \{\textit{of the probabilities on the path that take us from node } (1,1) \textit{ to node } (i-1,k)\}$$

$$\text{x}\ \frac{\partial C(i-1,k)}{\partial S(i,j)} e^{-(i-2)r_f \Delta t}$$

(21)

$$k = \begin{cases} j/2 & \textit{for even } j \\ (j+1)/2 & \textit{for odd } j \end{cases}$$

For example,

$$\frac{\partial C(1,1)}{\partial S(4,6)} = p(1,1)\left(1 - p(2,2)\right)\frac{\partial C(3,3)}{\partial S(4,6)} e^{-2r_f \Delta t}$$

$$\frac{\partial C(1,1)}{\partial S(5,3)} = \left(1 - p(1,1)\right)\left(1 - p(2,1)\right)p(3,1)\frac{\partial C(4,2)}{\partial S(5,3)} e^{-3r_f \Delta t}$$

V. Application using FTSE 100 options data

We use the daily closing prices of FTSE 100 call options of January 2003 to December 2003 as reported by LIFFE [9]. For the risk-free rate r_f, we use nonlinear cubic spline interpolation for matching each option contract with a continuous interest rate that corresponds to the option's maturity, by utilizing the 1-month to 12-month LIBOR offer rates, collected from Datastream.

Our initial sample (for the 12 months period) consists of 99051 observations. We adopt the following filtering rules:

[9] FTSE 100 options are traded with expiries in March, June, September, and December. Additional serial contracts are introduced so that options trade with expiries in each of the nearest 3 months. FTSE 100 options expire on the third Friday of the expiry month. FTSE 100 options positions are marked-to-market daily based on the daily settlement price, which is determined by LIFFE and confirmed by the Clearing House.

i) Eliminate calls for which the call price is greater than the value of the underlying asset, i.e. $C_{Mkt} > S(1,1)$. No observations are eliminated from this rule.

ii) Eliminate calls if the call price is less than its lower bound i.e. $C_{Mkt} < S(1,1)e^{-\delta T} - K\ e^{-r_f T}$. This rule eliminates 3206 observations.

iii) Eliminate calls with time to maturity less than 6 days, i.e. $T < 6$. This rule eliminates 3109 observations.

iv) Eliminate calls if their closing price is less than 0.5 index points. This rule eliminates 11373 observations.

v) Eliminate calls for which the trading volume is zero (since we want highly liquid options for calibration). This rule eliminates 66826 observations.

The final sample consists of 14537 observations.

In the implementation, for $\sigma(1)$ we use the at-the-money implied volatility given by LIFFE and for time to maturity, *T* we use the calendar days to maturity. Also, since the underlying asset of the options on FTSE 100 is a futures contract, we make the standard assumption that the dividend yield equals the risk free rate. The model is applied every day, with *n* = 6 and also with *n* = 7. For each implementation, the options used have the same underlying asset and the same time to maturity.

The evidence for the behaviour of the futures volatility in the literature is not clear. According to Samuelson (1965) the volatility of futures price changes should increase as the delivery date nears. However, Bessembinder et al. (1996) find that the Samuelson hypothesis is not supported for options on financials futures. In order to choose the value of λ that gives the best feasible solution we make consecutive draws from interval (5), which allows for both, positive and negative values of λ. The first value of λ is that of its lower bound. However, since dividend yield equals risk free rate, instead of $|r_f - \delta|$ we set 1.E-8. The next value of λ equals the old plus an appropriately chosen step size.

For brevity, we present results only for the first trading day of each month of the year 2003 and only for $n = 6$ (Table 1). *Trading Day* is the trading day of each contract, *Expiry* is the expiration month of each contract, *Asset* is the value of the underlying asset at the specified trading day, *N* is the number of contracts used for the calibration (the contracts that on the same trading day, have the same underlying asset and the same expiration day), *Error* is the value of the objective function, *Penalty* is the value of the penalty term. Ideally we want the error function and the penalty term to tend to zero. *Maturity* is the calendar days till the maturity of the contract, and *lambda* is the value of λ that gives the best feasible solution. Also, we present results only when the number of option contracts is greater than 3, since with fewer options the distribution of the underlying asset taken will not be reliable[10].

[Insert Table 1 here]

The results obtained support our modeling approach. As we can see in Table 1, in all cases the solution strictly satisfies the constraints since the penalty term equals zero. Also, we see that in 67 out of 69 cases, i.e. in 97.1% of the cases the error function tends to zero with an average value of 2.34E-08. In the other 2 cases, where the error function is greater than 1.E-4, the average error is 0.01. Similar results were found for $n = 7$.

Even though the problem requires a constrained non-convex optimization in $2(2^{n-1}-1)$ variables, the use of efficient optimization algorithms prevents the calibration of the model from becoming computationally too intensive. On average, the computational time in minutes required for each calibration had a mean (median) 1.10 (0.03) for $n = 6$ and 2.27 (0.08) for $n = 7$. The computer used for the calibration of the model had the following specifications: a Pentium 4 (3.2 GHz) CPU, Memory 1GB (RAM), and Windows XP Professional operating system. The codes were written in Matlab R2006a. The computational time needed would have decreased if the codes were written in the C/C++ language.

[10] In Table 1 we note that for the same contract (same underlying asset, same expiration) the number of contracts used in the model changes across months. That is because some contracts were removed because of the filtering rules.

When models provide an exact fit there is always the concern of over-fitting. We checked the model for over-fitting by pricing options with strikes in-between those used for the optimization (calibration). Then we made plots of the call prices (market prices and estimated from the model) versus moneyness. Over-fitting was also checked using a restricted sample consisting only of options with moneyness between 0.8 and 1.1, since these options are expected to be more liquid and more accurately priced[11]. For brevity, we exhibit only the plots for optimizations done in the first trading day of June (middle of the year) for the two samples using a tree with $n = 6$. As we see, for both samples the estimated call values increase smoothly with increasing moneyness without any evidence of over-fitting (see Fig.4). Similar results were obtained when a tree with $n = 7$ was used for the calibration procedure.

[Insert figure 4 here]

As a further check for over-fitting we use only part of the information to calibrate the tree and the other part to check the model using n = 6, 7, 8. Specifically, we leave out consecutively one of the N options at each time and we calibrate our model with the remaining options. In order to preserve the options' moneyness range stable and avoid problems of extrapolation, we do not remove the options with the highest and lowest moneyness. Over-fitting is checked like before using the full and the restricted sample of options. For the calibration only cases consisting of $N > 8$ were used. Results for the mean and median absolute errors are given in Table 2. We see that the error (given an average contract size of 90 for the full and 74.4 for the restricted sample) is small and rather stable[12].

[11] This sub-sample has a total of 13696 observations for the year 2003.

[12] Also, we compare our model (with respect to over-fitting) with the Black-Scholes model using the Whaley (1982) approach. According to this approach we find the volatility that minimizes the sum of square differences of the Black-Scholes option prices with their corresponding market prices using nonlinear minimization. Results show that the mean (median) absolute error using this approach is 7.36 (5.94) for the full sample and 6.61 (5.60) for the restricted sample which are much higher than the errors obtained using our model for n=6, 7, 8.

[Insert Table 2 here]

Since implied volatility changes with strike and time to maturity (volatility smile) the index should have a *non-lognormal* distribution which implies that the log-returns will deviate from normality. In order to see how realistic is the distribution obtained from our model for year 2003, we calculate the statistics of the 1-month log-returns obtained from our model and compare them with the historical 1-month log-returns for the year 2003 and the years 2001-2005. Specifically, for each calibration (with n = 6 and n = 7) for which the options maturity was between 28 and 32 calendar days, we calculate the first four moments (mean, variance, skewness and kurtosis). Then, in order to get a feeling for the representative statistics of 1-month log-returns we provide for each of those moments the mean and the median. The statistics for n = 6 are summarized in Table 3. Similar statistics were found for n = 7. Liu et al. (2005) discuss the derivations of historical, and implied real and risk-neutral distributions for the FTSE 100 index. They demonstrate that the needed adjustments to get the implied real variance, skewness and kurtosis from the implied risk-neutral ones are minimal. Thus, knowing that our implied risk-neutral moments (beyond the mean) are very close to the implied real ones, we can then compare them with the historical ones (without expecting the two distributions to be identical). As we would expect, the mean of the implied risk-neutral distribution of log-returns differs from that of the historical distribution. Also, as we see, both the implied risk-neutral and the historical distribution deviate from normality since they exhibit negative skewness and (mostly) excess kurtosis. This is an indication that the implied distribution is realistic.

[Insert Table 3 here]

In order to give further evidence for the implied distributions obtained by our model, representative implied distributions (histograms) for the 1-month log-returns in June 2003 are shown in Figures 5a (full sample) and 5b (restricted sample) for $n = 6$ and $n = 7$. To make the histograms of the implied distributions we make use of the Pearson system of distributions[13] as applied in Matlab[14]. Using the first four moments of the data it is easy to find in the Pearson system the distribution that matches these moments and to generate a random sample in order to produce a histogram corresponding to the implied distribution. From the figures, it is obvious that the implied distributions have negative skewness and positive kurtosis which is consistent with historical data. These figures are representative of the vast majority of cases[15]. Another interesting thing we observe is that distributions for $n = 6$ and $n = 7$ are practically indistinguishable for both samples.

[Insert Figures 5a, 5b here]

VI. Conclusions

In most options markets, the implied Black–Scholes volatilities vary with both strike and expiration, a relationship commonly known as the volatility smile. In this paper we capture the implied distribution from option market data using a non-recombining (binary) tree allowing the local volatility to be a function of the underlying asset and of time. The problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure, and use nonlinear constrained optimization to minimize the least squares error function on market prices. Specifically we adopt a penalty method and the optimization is implemented using a Quasi-Newton algorithm. Appropriate constraints allow us to maintain risk neutrality and to prevent arbitrage opportunities. The proposed model can accommodate European options with single maturities and, with minor modifications, options with multiple maturities. Also, this method is flexible since it applies to arbitrary underlying asset distributions, which implies arbitrary local volatility distributions. Market implied information embodied in the constructed tree

[13] In the Pearson system there is a family of distributions that includes a unique distribution corresponding to every valid combination of mean, standard deviation, skewness, and kurtosis.
[14] Copyright 2005 The MathWorks, Inc.
[15] In rare exceptions only we have implied distributions close to normal or even leptokurtic.

can help the pricing and hedging of exotic options and of OTC options on the same underlying process. We test our model using FTSE 100 options data. The results obtained strongly support our modelling approach. Pricing results are smooth without the presence of an over-fitting problem, and the derived implied distributions are realistic. Also, the computational burden is not a major issue.

APPENDIX 1A: Feasibility of the initialized non-recombining tree

We initialize the tree using the following volatility term structure:

$$\sigma(i) = \sigma(1)e^{\lambda(i-1)\Delta t} \quad , \lambda \in R \text{ where } i=1, \ldots, n$$

The feasibility of the initial tree depends on the right choice of the local volatility term structure; hence to obtain a feasible initial tree we must find an interval with the appropriate values of λ. In order to preserve the risk neutrality at every time step, the following constraints must be satisfied:

$$S(i,j)e^{(r_f-\delta)\Delta t} \le S(i+1,2j) \quad \text{(A 1a)}$$

$$S(i,j)e^{(r_f-\delta)\Delta t} \ge S(i+1,2j-1) \quad \text{(A 1b)}$$

Also,

$$S(i+1,2j) = S(i,j)u(i,j) = S(i,j)e^{\sigma(i)\sqrt{\Delta t}} \quad \text{(A 2a)}$$

$$S(i+1,2j-1) = S(i,j)d(i,j) = S(i,j)e^{-\sigma(i)\sqrt{\Delta t}} \quad \text{(A 2b)}$$

Substituting (A 2a) and (A 2b) to (A 1a) and (A 1b) respectively we get the following inequalities:

$$\sigma(i) \ge (r_f - \delta)\sqrt{\Delta t} \quad \text{(A 3a)}$$

$$\sigma(i) \ge -(r_f - \delta)\sqrt{\Delta t} \quad \text{(A 3b)}$$

Thus we have that

$$\sigma(i) \ge \left| r_f - \delta \right| \sqrt{\Delta t} \qquad \forall i \quad \text{(A 4)}$$

For $\lambda \ge 0$, $\sigma(i) = \sigma(1)e^{\lambda(i-1)\Delta t}$ is strictly increasing.

Since (A 4) holds for every i this means that

$$\min \sigma(i) \ge \left| r_f - \delta \right| \sqrt{\Delta t} \qquad \text{or}$$

$$\sigma(1) \ge \left| r_f - \delta \right| \sqrt{\Delta t} \quad \text{(A 5)}$$

The minimum value of $\sigma(i)$ is for $i=1$ ($\sigma(1)$), thus (A 5) is independent of λ. Therefore, if λ is positive there is no upper bound for λ.

For $\lambda < 0$, $\sigma(i) = \sigma(1)e^{\lambda(i-1)\Delta t}$ is strictly decreasing.

Since (A 4) holds for every *i* this means that

$$\min \sigma(i) \geq \left| r_f - \delta \right| \sqrt{\Delta t} \Rightarrow$$

$$\sigma(n) \geq \left| r_f - \delta \right| \sqrt{\Delta t} \Rightarrow$$

$$e^{\lambda(n-1)\Delta t} \geq \frac{\left| r_f - \delta \right| \sqrt{\Delta t}}{\sigma(1)}$$

But, $(n-1)\Delta t = T$, thus,

$$\lambda \geq \frac{1}{T} \log\left(\frac{\left| r_f - \delta \right| \sqrt{\Delta t}}{\sigma(1)} \right) \qquad \text{(A 6)}$$

If we allow λ to take both negative and positive values, then λ should belong in the interval,

$$\lambda \in \left[\frac{1}{T} \log\left(\frac{\left| r_f - \delta \right| \sqrt{\Delta t}}{\sigma(1)} \right), \quad +\infty \right) \qquad \text{(A 7)}$$

APPENDIX 1B: Feasibility of the initialized non-recombining tree assuming time dependent r_f, δ and Δt

We denote with $r_f(i)$ and $\delta(i)$ the risk free rate and dividend yield respectively between two consecutive time steps, i.e. between time step i and $i+1$, $i = 1, ..., n-1$.

[Insert Figure A1]

We initialize the tree using the following volatility term structure:

$$\sigma(i) = \sigma(1)e^{\lambda\sum_{j=1}^{i-1}\Delta t(j)} \quad , \ \lambda \in R \text{ where } i=1, \ldots, n$$

The feasibility of the initial tree depends on the right choice of the local volatility term structure; hence to obtain a feasible initial tree we must find an interval with the appropriate values of λ. In order to preserve the risk neutrality at every time step, the following constraints must be satisfied:

$$S(i,j)e^{(r_f(i)-\delta(i))\Delta t(i)} \leq S(i+1, 2j) \qquad \text{(A 1a′)}$$

$$S(i,j)e^{(r_f(i)-\delta(i))\Delta t(i)} \geq S(i+1, 2j-1) \qquad \text{(A 1b′)}$$

Also,

$$S(i+1, 2j) = S(i,j)u(i,j) = S(i,j)e^{\sigma(i)\sqrt{\Delta t}(i)} \qquad \text{(A 2a′)}$$

$$S(i+1, 2j-1) = S(i,j)d(i,j) = S(i,j)e^{-\sigma(i)\sqrt{\Delta t}(i)} \qquad \text{(A 2b′)}$$

Substituting (A 2a′) and (A 2b′) to (A 1a′) and (A 1b′) respectively we get the following inequalities:

$$\sigma(i) \geq (r_f(i) - \delta(i))\sqrt{\Delta t(i)} \qquad \text{(A 3a′)}$$

$$\sigma(i) \geq -(r_f(i) - \delta(i))\sqrt{\Delta t(i)} \qquad \text{(A 3b′)}$$

Thus we have that

$$\sigma(i) \geq \left|r_f(i) - \delta(i)\right|\sqrt{\Delta t(i)} \qquad \forall i \qquad \text{(A 4′)}$$

For $\lambda \geq 0$, $\sigma(i) = \sigma(1)e^{\lambda\sum_{j=1}^{i-1}\Delta t(j)}$ is strictly increasing.

Let $\xi_M = \max_i \left|r_f(i) - \delta(i)\right|\sqrt{\Delta t(i)}$

Then (A 4′) holds for every i if

$$\min_i \sigma(i) \geq \xi_M \qquad \text{or}$$

$$\sigma(1) \geq \xi_M \tag{A 5′}$$

The minimum value of $\sigma(i)$ is for $i = 1$ $(\sigma(1))$, thus (A 5′) is independent of λ. Therefore, if λ is positive there is no upper bound for λ.

For $\lambda < 0$, $\sigma(i) = \sigma(1)e^{\lambda \sum_{j=1}^{i-1} \Delta t(j)}$ is strictly decreasing.

Let $\xi_m = \min_i \left| r_f(i) - \delta(i) \right| \sqrt{\Delta t(i)}$

Then (A 4′) holds for every i if

$$\min_i \sigma(i) \geq \xi_m$$

$$\sigma(n) \geq \xi_m$$

$$\sigma(1)e^{\lambda \sum_{j=1}^{n-1} \Delta t(j)} \geq \xi_m$$

But, $\sum_{j=1}^{n-1} \Delta t(j) = T$, thus,

$$\lambda \geq \frac{1}{T} \log\left(\frac{\xi_m}{\sigma(1)} \right) \tag{A 6′}$$

If we allow λ to take both negative and positive values, then λ should belong in the interval,

$$\lambda \in \left[\frac{1}{T} \log\left(\frac{\xi_m}{\sigma(1)} \right), \quad +\infty \right) \tag{A 7′}$$

Appendix 2: Formulas adjusted for time dependent r_f, δ and Δt

We denote with $r_f(i)$ and $\delta(i)$ the risk free rate and dividend yield respectively between two consecutive time steps, i.e. between time step i and $i+1$, $i = 1, ..., n-1$ and with r'_f and δ' we denote the risk free rate and dividend yield respectively from today till the maturity of the option, i.e. from $i = 1$ to $i = n$.

If we allow r_f, δ and Δt to be time dependent the equations of the main text are replaced with the following:

$$u(i,j) = e^{\sigma(i)\sqrt{\Delta t(i)}} \tag{3a'}$$

$$d(i,j) = e^{-\sigma(i)\sqrt{\Delta t(i)}} = \frac{1}{u(i,j)}, \quad i = 1, \ ..., \ n-1 \ , \quad j = 1, \ ..., \ 2^{i-1} \tag{3b'}$$

$$\lambda \in \left[\frac{1}{T} \log\left(\frac{\xi_m}{\sigma(1)} \right), \quad +\infty \right) \tag{5'}$$

where,

$$\xi_m = \min_i \left| r_f(i) - \delta(i) \right| \sqrt{\Delta t(i)}$$

$$p(i,j) = \frac{S(i,j) e^{(r_f(i)-\delta(i))\Delta t(i)} - S(i+1, 2j-1)}{S(i+1, 2j) - S(i+1, 2j-1)}, \ i = 1, \ ..., \ n-1 \ , \quad j = 1, \ ..., \ 2^{i-1} \tag{7'}$$

$$C(i,j) = \left(p(i,j) C(i+1, 2j) + (1 - p(i,j)) C(i+1, 2j-1) \right) e^{-r_f(i)\Delta t(i)} \tag{9b'}$$

$$i = n-1, ..., 1, \quad j = 1, ..., \ 2^{i-1}$$

$$S(i,j) e^{(r_f(i)-\delta(i))\Delta t(i)} \le S(i+1, 2j), \quad i = 1, \ ..., n-1, \quad j = 1, ..., \ 2^{i-1} \tag{10a'}$$

$$S(i,j) e^{(r_f(i)-\delta(i))\Delta t(i)} \ge S(i+1, 2j-1), \quad i = 1, \ ..., n-1, \quad j = 1, ..., \ 2^{i-1} \tag{10b'}$$

$$\max\left(S(1,1) e^{-\delta' T} - K e^{-r'_f T}, 0 \right) \le C_{Mod} \le S(1,1) \tag{11'}$$

$$g_1(i,j) = S(i,j)e^{(r_f(i)-\delta(i))\Delta t(i)} - S(i+1,2j-1) \geq 0,\ i = 1, ..., n-1,\ j = 1, ..., 2^{i-1} \quad (14a')$$

$$g_2(i,j) = S(i+1,2j) - S(i,j)e^{(r_f(i)-\delta(i))\Delta t(i)} \geq 0,\ i = 1, ..., n-1,\ j = 1, ..., 2^{i-1} \quad (14b')$$

$$g_3(k) = S(1,1) - C_{Mod}(k) \geq 0,\ \ k = 1, ..., N \quad (14c')$$

$$g_4(k) = C_{Mod}(k) - \max(S(1,1)e^{-\delta' T} - K(k)e^{-r_f' T}, 0) \geq 0,\ \ k = 1, ..., N \quad (14d')$$

$$\nabla_{S_{i,j}^{(l)}} p(i,j) \equiv \begin{bmatrix} \partial p(i,j) / \partial S(i,j) \\ \partial p(i,j) / \partial S(i+1,2j) \\ \partial p(i,j) / \partial S(i+1,2j-1) \end{bmatrix} = \frac{1}{S(i+1,2j) - S(i+1,2j-1)} \begin{bmatrix} e^{(r_f(i)-\delta(i))\Delta t(i)} \\ -p(i,j) \\ -(1-p(i,j)) \end{bmatrix} \quad (17')$$

$$\nabla_{S_{i,j}^{(l)}} C(i,j) \equiv \begin{bmatrix} \partial C(i,j) / \partial S(i,j) \\ \partial C(i,j) / \partial S(i+1,2j) \\ \partial C(i,j) / \partial S(i+1,2j-1) \end{bmatrix} = \begin{bmatrix} \Delta(i,j) \\ p(i,j)\left(\Delta(i+1,2j) - \Delta(i,j)e^{\delta(i)\Delta t(i)}\right)e^{-r_f(i)\Delta t(i)} \\ \left(1-p(i,j)\right)\left(\Delta(i+1,2j-1) - \Delta(i,j)e^{\delta(i)\Delta t(i)}\right)e^{-r_f(i)\Delta t(i)} \end{bmatrix}$$

(18′)

$$\Delta(i,j) = \frac{C(i+1,2j) - C(i+1,2j-1)}{S(i+1,2j) - S(i+1,2j-1)} e^{-\delta(i)\Delta t(i)} = \frac{\partial C(i,j)}{\partial S(i,j)} \equiv Delta\ \ Ratio \quad (19')$$

$$\frac{\partial C(1,1)}{\partial S(i,j)} = \prod \{of\ the\ probabilities\ on\ the\ path\ that\ take\ us\ from\ node\ (1,1)\ to\ node\ (i-1,k)\}$$

$$\text{x}\ \frac{\partial C(i-1,k)}{\partial S(i,j)} e^{-\sum_{h=2}^{i-1} r_f(h)\Delta t(h)}$$

(21′)

$$k = \begin{cases} j/2 & for\ \ even\ \ j \\ (j+1)/2 & for\ \ odd\ \ j \end{cases}$$

References

Ait-Sahalia, Y., Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance, 53*, pp. 499-547.

Barle, S., Cakici, N. (1995). Growing a smiling tree. *Risk*, 8, pp. 76-81.

Bessembinder, H., Coughenour, J. F., Seguin, P. J. (1996). Is there a term structure of futures volatilities? Re-evaluating the Samuelson Hypothesis. *Working Paper*.

Black, F., and Scholes, M. (1973). Pricing of options and corporate liabilities. *Journal of Political Economy*, 81, pp. 637-659.

Britten-Jones, M., and Neuberger, A. (2000). Option prices, implied price processes and stochastic volatility. *Journal of Finance,* 55, pp. 839-866.

Charalambous, C. (1992). Conjugate gradient algorithm for efficient training of artificial neural networks. *IEE Proceedings-G*, 139, pp. 301-310.

Chriss, N. (1996). Transatlantic trees. *Risk,* 9, pp. 143-159.

Cox, J., Ross, S., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7, pp. 229-263.

Derman, E., Kani, I. (1994). Riding on a smile. *Risk*, 7(2), pp. 32-39.

Derman, E., Kani, I., and Chriss, N. (1996). Implied trinomial trees of the volatility smile. *Journal of Derivatives*, 3, pp. 7-22.

Derman, E., Kani, I. (1998). Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility. *International Journal of Theoretical and Applied Finance*, 1, pp. 61-110.

Dupire, B. (1994). Pricing with a smile. *Risk*, 7, pp. 18-20.

Edirisinghe, Naik, and Uppal, (1993). Optimal replication of options with transaction costs and trading restrictions. *Journal of Financial and Quantitative Analysis,* 28, pp. 117-138.

Fiacco, A. V., McCormick, G. P. (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley and Sons, Inc.

Fletcher, R. (1987). *Practical methods of optimization.* Chichester, Wiley.

Jackwerth, J., and Rubinstein, M. (1996). Recovering probability distributions from option prices. *Journal of Finance*, 51, pp. 1611-1631.

Jackwerth, J. (1997). Generalized binomial trees. *Journal of Derivatives*, 5, pp. 7-17.

Ledoit, O., Santa-Clara, P. (1998). Relative pricing of options with stochastic volatility. *Working Paper*, Univ. of California, Los Angeles.

Liu, X., Shackleton, B. M., Taylor, J. S., Xu, X. (2005). Closed transformations from risk-neutral to real-world distributions. *Working Paper*, Univ. of Lancaster.

Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance,* 49, pp. 771-818.

Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6, pp. 41-49.

Skiadopoulos, G. (2001). Volatility smile consistent option models: A survey. *International Journal of Theoretical and Applied Finance*, 4, pp. 403-437.

Stutzer, M. (1996). A simple nonparametric approach to derivative security valuation. *Journal of Finance,* 51, pp. 1633-1652.

Talias, K. (2005). Implied binomial trees and genetic algorithms. *Doctoral Dessertation,* Imperial College, London.

Whaley, R. (1982). Valuation of American call options on dividend-paying stocks. *Journal of Financial Economics*, 10, pp. 29-58.

FIGURES

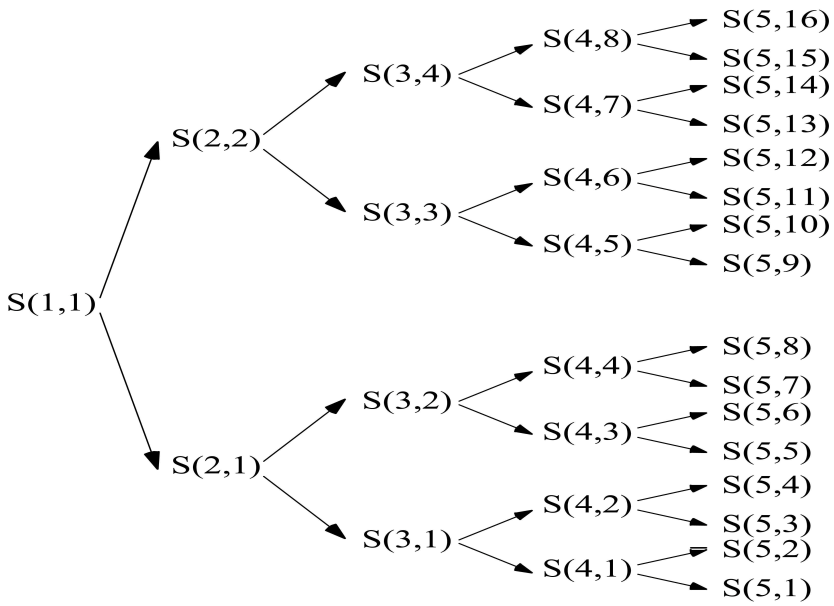


Figure 1: Non-recombining tree with 4 steps.

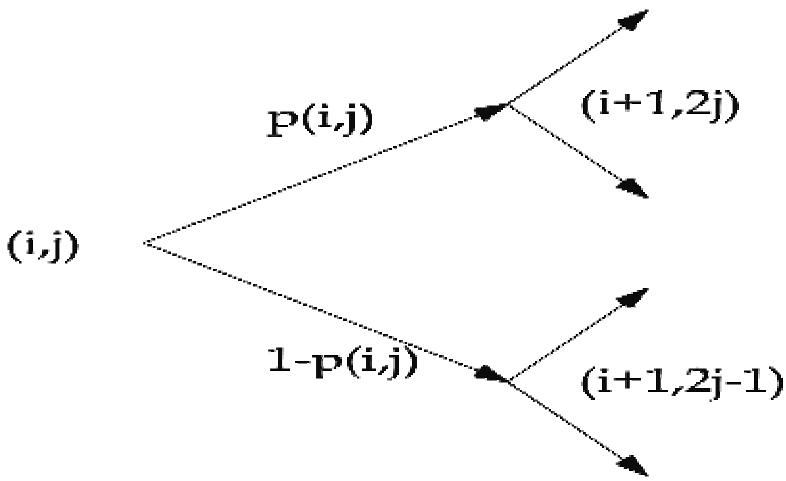


Figure 2: A typical triplet in a non-recombining tree.

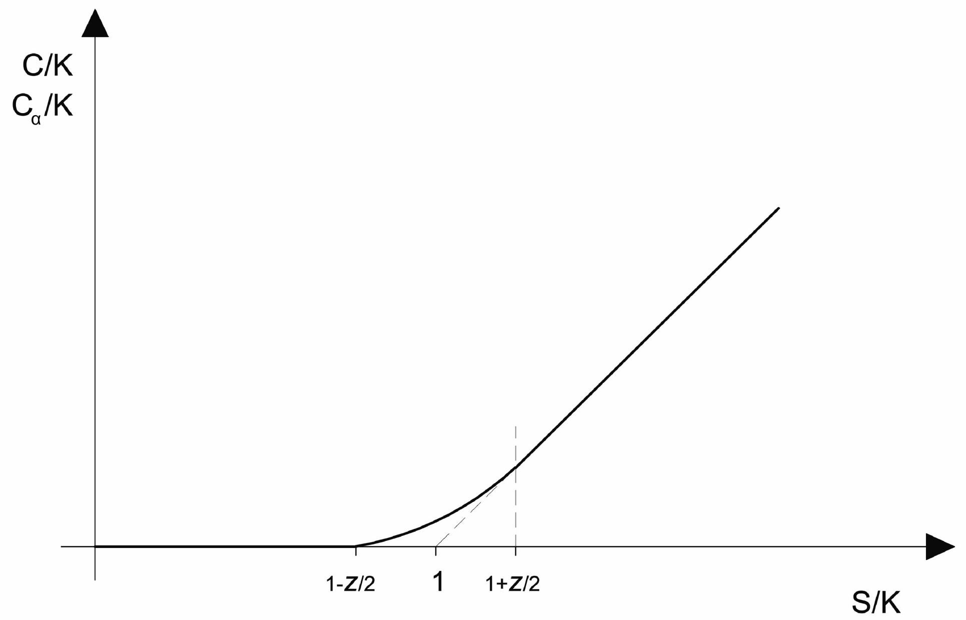
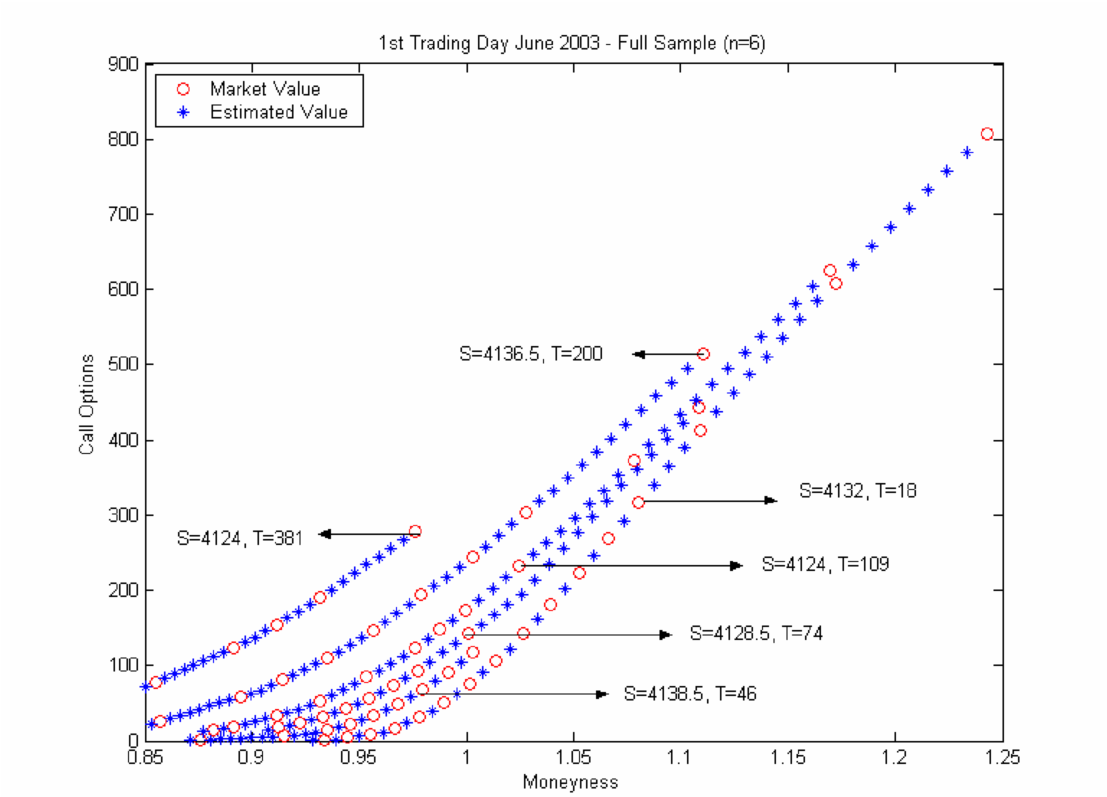


Figure 3: Smoothing of the option pay-off function at maturity.



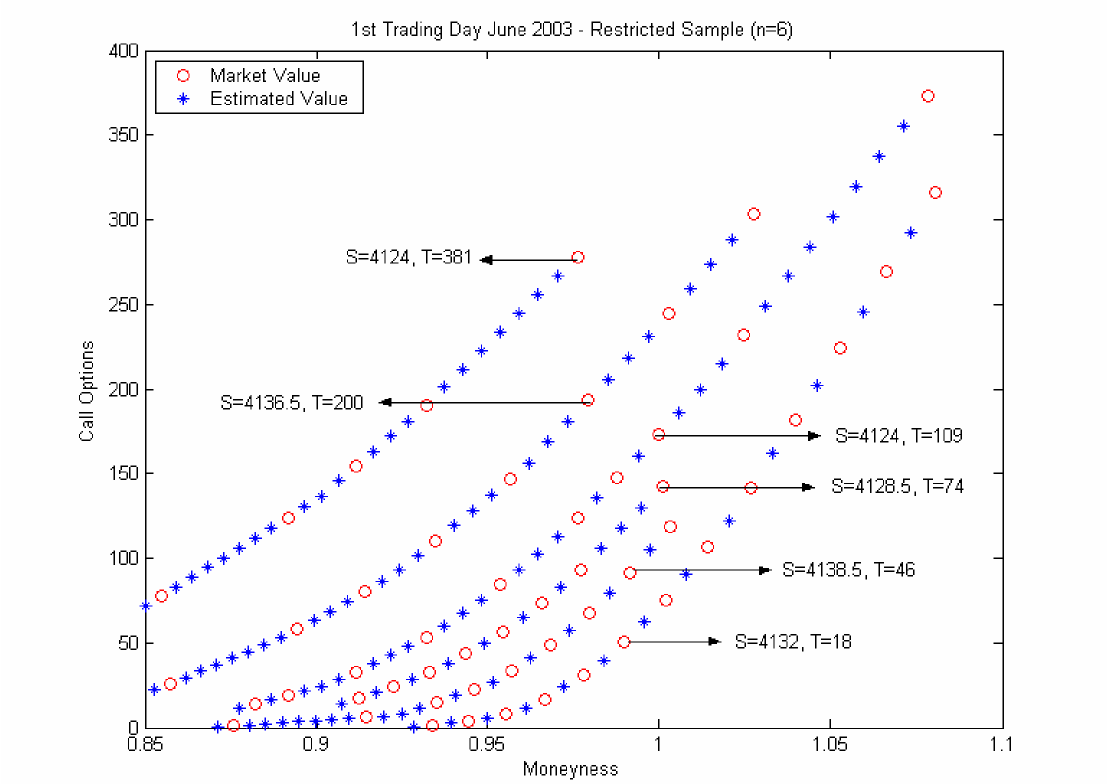


Figure 4: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003. *S* denotes the value of the underlying asset and *T* the calendar days to maturity.

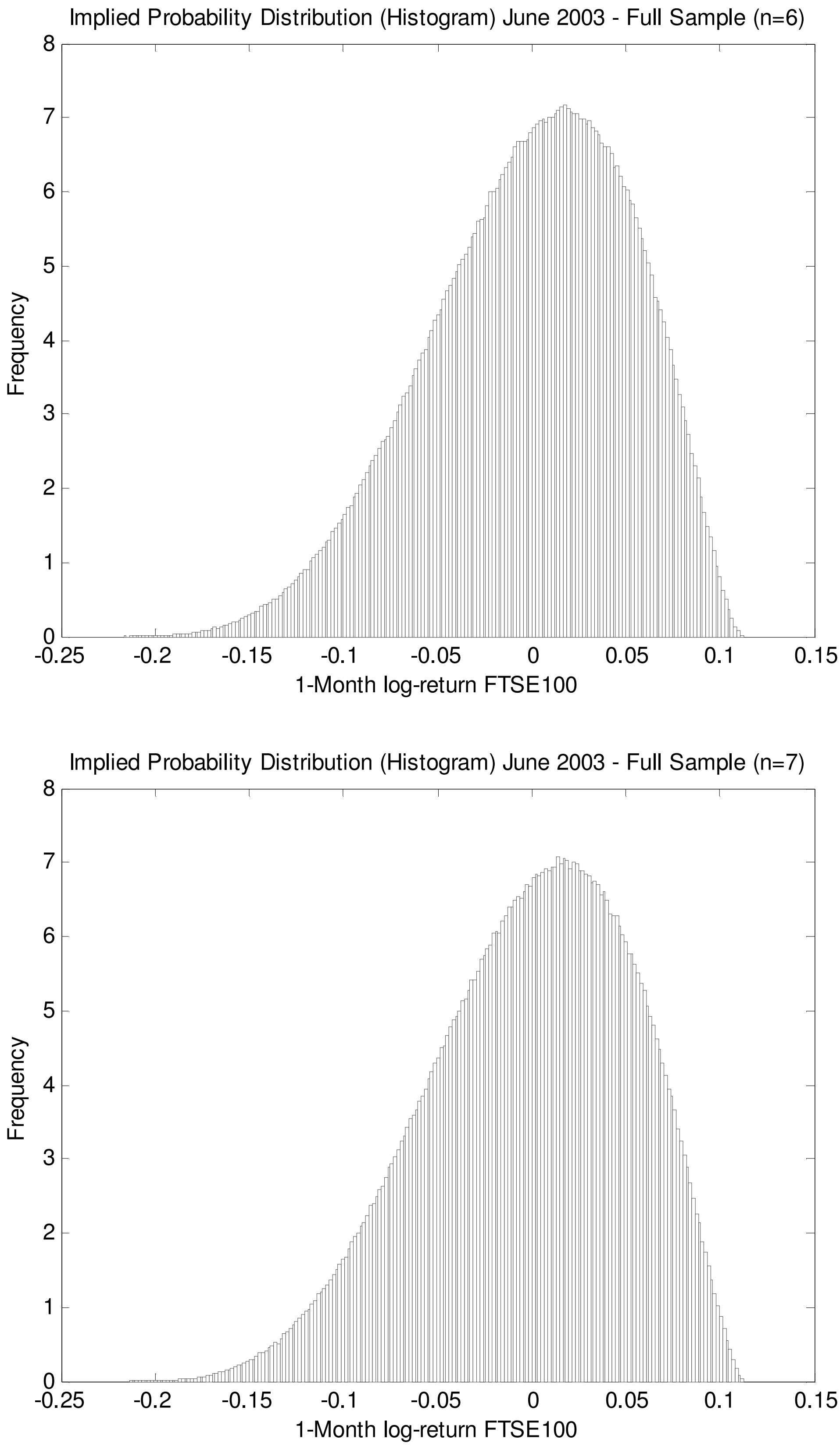
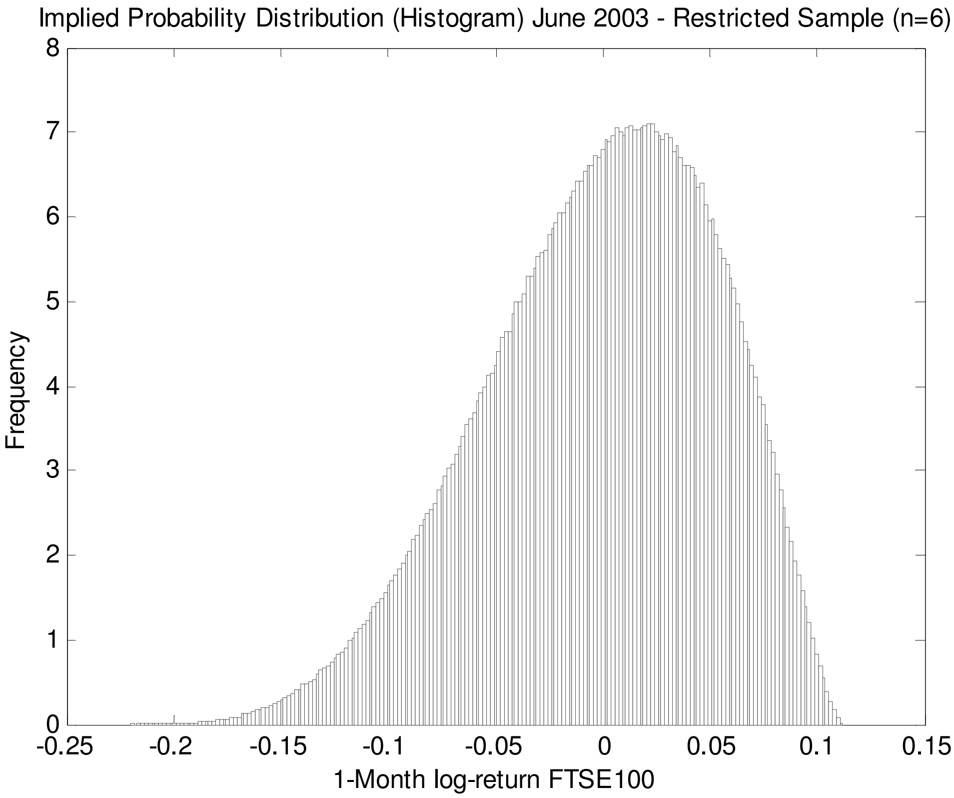


Figure 5a: Implied probability distributions (histograms) obtained for the 1-month log-return of June 2003 using the full sample for *n* = *6* and *n* = *7*.



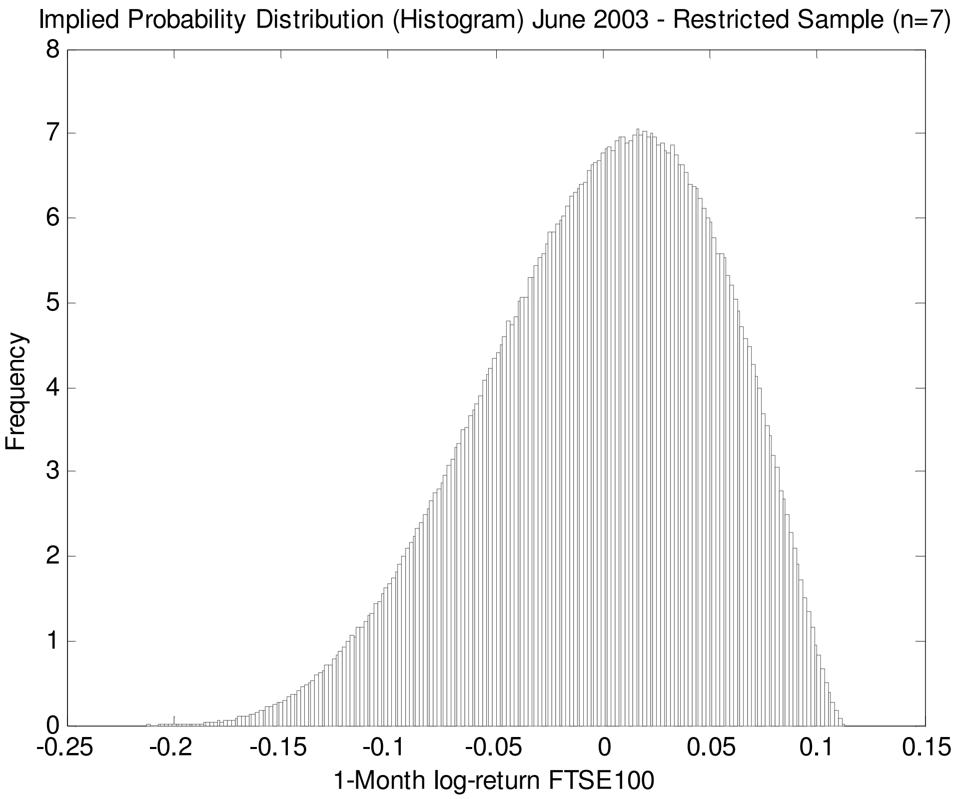


Figure 5b: Implied probability distributions (histograms) obtained for the 1-month log-return of June 2003 using the restricted sample for $n = 6$ and $n = 7$.

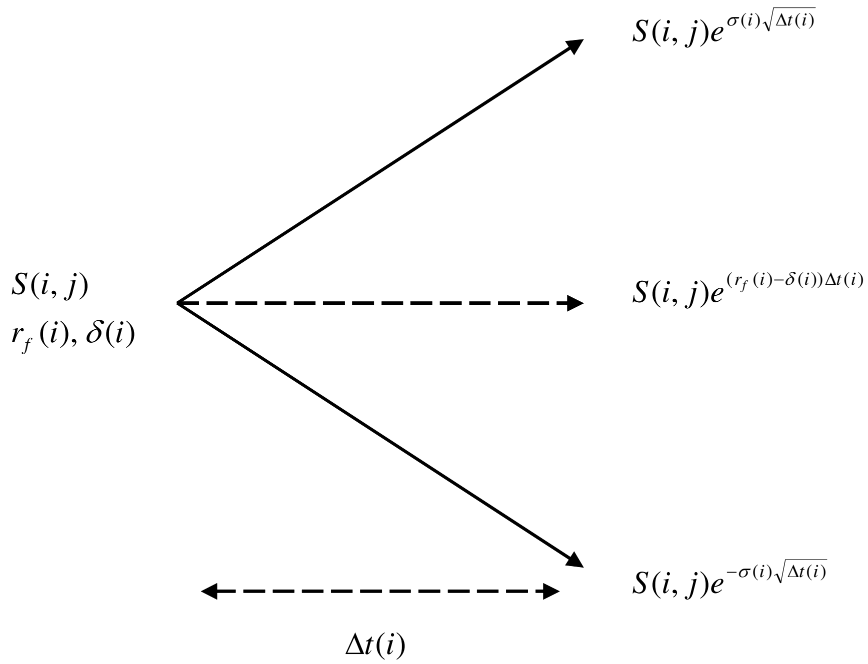


Figure A1: A typical triplet in the initialization of the non-recombining tree assuming r_f, δ and Δt to be time dependent.

TradingDay	*Expiry*	*Asset*	*N*	*Error*	*Penalty*	*Maturity*	*lambda*
01/02/2003	Jan-03	4014	17	7.933E-11	0	15	4.3429
01/02/2003	Feb-03	4019	19	4.2E-05	0	50	-1.3851
01/02/2003	Mar-03	3991	12	5.855E-12	0	78	-6.6823
01/02/2003	Jun-03	3995	11	2.721E-13	0	169	-0.3963
01/02/2003	Dec-03	3999	6	0.0208333	0	351	0.4096
02/03/2003	Feb-03	3675.5	16	7.254E-08	0	18	-3.9616
02/03/2003	Mar-03	3646	14	8.611E-11	0	46	-1.5180
02/03/2003	Apr-03	3644.5	15	2.722E-12	0	73	-6.4346
02/03/2003	May-03	3645	6	2.355E-14	0	102	-6.6918
02/03/2003	Jun-03	3647	7	2.196E-12	0	137	-4.9425
02/03/2003	Sep-03	3640	5	1.337E-14	0	228	-1.1252
02/03/2003	Dec-03	3653.5	7	5.859E-11	0	319	-0.9666
03/03/2003	Mar-03	3657	16	6.572E-13	0	18	-3.9616
03/03/2003	Apr-03	3655	13	2.573E-11	0	45	-1.5455
03/03/2003	May-03	3655	9	2.466E-12	0	74	-8.3274
03/03/2003	Jun-03	3655.5	7	9.825E-14	0	109	-6.2312
03/03/2003	Sep-03	3645	9	1.735E-12	0	200	-3.3359
04/01/2003	Apr-03	3684.5	16	4.548E-11	0	16	5.0218
04/01/2003	May-03	3683.5	16	4.396E-12	0	45	-1.5510
04/01/2003	Jun-03	3686.5	10	2.096E-11	0	80	-5.0920
04/01/2003	Jul-03	3693	7	1.222E-11	0	108	-6.1873
04/01/2003	Sep-03	3676.5	5	4.563E-11	0	171	-3.9295
04/01/2003	Mar-04	3667	5	1.301E-11	0	352	-1.8670

Table 1: Results for the application of the model on the 1st trading day of each month of the year 2003: *Trading Day* is the trading day of each contract, *Expiry* is the expiration month of each contract, *Asset* is the value of the underlying asset at the specified trading day, *N* is the number of contracts used for the calibration, *Error* is the value of the objective function, *Penalty* is the value of the penalty term, *Maturity* is the calendar days till the maturity of the contract and *lambda* is the value of λ that gives the best feasible solution

TradingDay	*Expiry*	*Asset*	*N*	*Error*	*Penalty*	*Maturity*	*lambda*
05/01/2003	May-03	3874	15	8.178E-08	0	15	-4.1763
05/01/2003	Jun-03	3879	14	1.581E-11	0	50	-1.3588
05/01/2003	Jul-03	3885.5	10	2.535E-11	0	78	-8.6239
05/01/2003	Sep-03	3870.5	4	2.657E-12	0	141	-2.1062
05/01/2003	Mar-04	3869	5	2.448E-13	0	322	-1.8172
06/02/2003	Jun-03	4132	16	3.535E-09	0	18	25.1320
06/02/2003	Jul-03	4138.5	9	6.685E-12	0	46	-1.4751
06/02/2003	Aug-03	4128.5	9	9.668E-13	0	74	-9.0523
06/02/2003	Sep-03	4124	11	5.437E-11	0	109	-6.0923
06/02/2003	Dec-03	4136.5	9	3.798E-14	0	200	-3.2689
06/02/2003	Jun-04	4124	5	1.825E-11	0	381	-1.6875
07/01/2003	Jul-03	3967	13	0.0007343	0	17	13.5874
07/01/2003	Aug-03	3958	12	4.422E-12	0	45	0.1871
07/01/2003	Sep-03	3955	12	3.807E-11	0	80	-1.5851
07/01/2003	Oct-03	3959	4	5.92E-14	0	108	-6.1578
07/01/2003	Dec-03	3964	11	1.544E-12	0	171	-1.8451
07/01/2003	Mar-04	3956	7	1.265E-13	0	261	-2.4906
08/01/2003	Aug-03	4091.5	11	7.597E-12	0	14	1.1396
08/01/2003	Sep-03	4088.5	14	7.022E-11	0	49	5.6305
08/01/2003	Oct-03	4094.5	4	5.318E-11	0	77	-8.6243
08/01/2003	Nov-03	4096.5	5	2.43E-13	0	112	-5.8741
08/01/2003	Dec-03	4100.5	5	1.804E-14	0	140	-0.4681
08/01/2003	Mar-04	4097	4	1.013E-12	0	230	-1.7231

Table 1 (continued): Results for the application of the model on the 1st trading day of each month of the year 2003: *Trading Day* is the trading day of each contract, *Expiry* is the expiration month of each contract, *Asset* is the value of the underlying asset at the specified trading day, *N* is the number of contracts used for the calibration, *Error* is the value of the objective function, *Penalty* is the value of the penalty term, *Maturity* is the calendar days till the maturity of the contract and *lambda* is the value of λ that gives the best feasible solution.

TradingDay	*Expiry*	*Asset*	*N*	*Error*	*Penalty*	*Maturity*	*lambda*
08/01/2003	Jun-04	4098.5	6	7.7014E-12	0	321	0.9996
09/01/2003	Sep-03	4215	16	9.231E-11	0	18	-3.8161
09/01/2003	Oct-03	4222	9	4.986E-13	0	46	-1.4569
09/01/2003	Nov-03	4225	9	4.4323E-12	0	81	-8.1673
09/01/2003	Dec-03	4229	12	1.0014E-12	0	109	-0.6026
09/01/2003	Mar-04	4224.5	4	1.1111E-11	0	199	-32.5850
10/01/2003	Oct-03	4162.5	12	1.1715E-06	0	16	-4.3573
10/01/2003	Nov-03	4167	12	1.7641E-13	0	51	-1.3282
10/01/2003	Dec-03	4169.5	19	2.796E-12	0	79	-8.4832
10/01/2003	Jan-04	4173.5	5	5.4217E-13	0	107	-62.1642
10/01/2003	Mar-04	4162	8	1.1045E-11	0	169	-3.8885
10/01/2003	Jun-04	4171.5	5	1.0698E-13	0	260	-2.4977
11/03/2003	Nov-03	4330	12	9.9737E-12	0	18	26.2023
11/03/2003	Dec-03	4333	15	4.1511E-11	0	46	0.5538
11/03/2003	Jan-04	4344	9	3.2388E-12	0	74	-8.7229
11/03/2003	Feb-04	4354	7	2.0997E-12	0	109	-2.1552
11/03/2003	Mar-04	4329	7	2.6622E-13	0	136	-4.7931
11/03/2003	Jun-04	4343.5	7	4.4587E-13	0	227	-28.4200
12/01/2003	Dec-03	4415.5	13	6.9791E-11	0	18	15.2128
12/01/2003	Jan-04	4426	10	2.5172E-13	0	46	-1.4475
12/01/2003	Feb-04	4433.5	13	3.7691E-12	0	81	-8.1270
12/01/2003	Mar-04	4410.5	10	5.3369E-12	0	108	-6.0614
12/01/2003	Jun-04	4423.5	4	9.9101E-16	0	199	-3.2410

Table 1 (continued): Results for the application of the model on the 1st trading day of each month of the year 2003: *Trading Day* is the trading day of each contract, *Expiry* is the expiration month of each contract, *Asset* is the value of the underlying asset at the specified trading day, *N* is the number of contracts used for the calibration, *Error* is the value of the objective function, *Penalty* is the value of the penalty term, *Maturity* is the calendar days till the maturity of the contract and *lambda* is the value of λ that gives the best feasible solution.

Model		Absolute Errors	
		Full Sample	Restricted Sample
n =6	Mean	1.2458	1.1065
	Median	0.9163	0.8709
n =7	Mean	1.1375	0.9792
	Median	0.8005	0.6929
n =8	Mean	1.1286	0.9350
	Median	0.7928	0.6886
	Observations	446	405

Table 2: Mean and median absolute errors using our model for *n*=6, 7, 8 and data from the full and the restricted sample.

Implied (2003, n = 6)	Mean	Variance	Skewness	Kurtosis	Observations
Mean	-0.0024	0.0048	-0.6938	4.5075	58
Median	-0.0013	0.0027	-0.6653	3.6405	58
Historical	**Mean**	**Variance**	**Skewness**	**Kurtosis**	**Observations**
2003	0.0106	0.0014	-0.6572	2.7689	12
2001-2005	-0.0021	0.0018	-1.1177	4.4749	59

Table 3: Implied risk-neutral and historical statistics of the distribution of the FTSE 100 1-month log-returns